DOCUMENT RESUME

ED 169 130 TM 008 679

TITLE	Effects of Coaching on Scholastic Aptitude Test Scores.
INSTITUTION	College Entrance Examination Board, New York, N.Y.
PUB DATE	68
NOTE	29p.
AVAILABLE FROM	College Entrance Examination Board, Publications Order Office, Box 592, Princeton, New Jersey 08540 ($0.25)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Academic Aptitude; *Achievement Gains; Aptitude Tests; *College Bound Students; *College Entrance Examinations; *Scores; Senior High Schools; Testing Problems; Test Reliability; *Test Wiseness
IDENTIFIERS	*Scholastic Aptitude Test

ABSTRACT

Intended primarily for secondary school administrators, teachers, and counselors, this booklet summarizes the information available, as of 1965, about the ways special coaching or tutoring may affect students' scores on the College Board's Scholastic Aptitude Test (SAT). Coaching refers to the variety of methods used in attempting to increase, in a relatively short time, students' mastery of the particular skills, concepts, and reasoning abilities tested by the SAT. Seven studies of the effects of coaching on SAT scores are summarized; four of these studies were conducted by Educational Testing Service for the College Board. The subjects varied in ability and academic background; they included public, private, and vocational high school students. The amount of coaching, instructional materials used, and coaching experience of tutors also varied. The net result across all studies was that score gains directly attributable to coaching were, on the average, fewer than ten points--a difference of such small magnitude on the SAT score scale of 200 to 800 that the College Board considered it unreasonable to expect it to affect college admissions decisions. The magnitude of the gains resulting from coaching were always small, regardless of the coaching method used or the differences in the students coached. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Effects of Coaching on Scholastic Aptitude Test Scores

ED169130

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

S. W. Gardner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM008679

One complimentary copy of this booklet has been supplied to each secondary school; a maximum of five additional complimentary copies may be obtained by secondary schools on request. For all other requests, copies are available at 25 cents per copy from the College Entrance Examination Board, Publications Order Office, Box 592, Princeton, New Jersey 08540.

Editorial inquiries regarding this publication should be addressed to Editorial Office, College Entrance Examination Board, 475 Riverside Drive, New York, New York 10027.

Copyright 1965, 1968 by College Entrance Examination Board, New York
Library of Congress Catalog Number: 65-27938

Printed in the United States of America

Contents

Summary of research findings on coaching 4

Changes in a student's Scholastic Aptitude Test scores between junior and senior years 5

Statement by College Board trustees 8

Seven studies of coaching 12

Coaching boys of above-average ability in independent preparatory schools (Dyer, 1953) 12

Coaching boys and girls in public schools (French, 1955) . . 14

Coaching students by using actual items from the SAT they later take (French, 1955) 18

Intensive coaching of small groups of students (Dear, 1958) . 18

Commercial coaching programs: their effect on public-school boys and girls of above-average ability (Frankel, 1960) . . . 21

Intensive tutoring of public-school boys and girls in vocabulary, reading skills, concept mastery, problem solving (Whitla, 1962) 23

Informal review of test scores of coached and uncoached students within one public school (Lass, 1958) 25

Sources 28

Summary of research findings on coaching

The purpose of this booklet is to summarize the information now available about the ways special coaching or tutoring may affect students' scores on the College Board's Scholastic Aptitude Test. It is intended primarily for secondary school administrators, teachers, and counselors—all those who determine policy with regard to secondary school curriculum and services and those who teach and counsel prospective college students.

The word coaching, as used here, refers to a variety of methods used in attempting to increase in a relatively short time students' mastery of the particular skills, concepts, and reasoning abilities tested by the SAT. Many schools have been importuned by students and parents to provide such coaching, and some have tried to do so. Increasing numbers of commercial entrepreneurs have been advertising the promise of quickly attained increases in SAT scores through programs of practice and drill.

Respecting the motives of students who seriously want to improve their scholastic aptitude, and of parents and educators who want to assist them, the College Board has explored at length the question of whether or not coaching does in fact achieve its stated purpose of improving students' performance on the SAT.

Despite variable factors from one study to another, the net result across all studies is that score gains directly attributable to coaching amount, on the average, to fewer than 10 points—a difference of such small magnitude on the SAT score scale of 200 to 800 that it is unreasonable to expect it to affect college admissions decisions. The magnitude of the gains resulting from coaching vary slightly, but they are always small regardless of the coaching method used or the differences in the students coached.

The seven studies on which this conclusion is based are reported in the pages that follow. Four of these studies were conducted by Educational Testing Service for the College Board, and three were carried

out by independent investigators. There were certain common elements in all the studies: First, before each study, the students involved took the SAT. Second, some of the students were then exposed to coaching of some kind, while others received no coaching. Third, a different edition of the SAT was administered to both coached and uncoached students 7 to 10 months after they took the test the first time. Fourth, comparisons were then made to see whether the coached students showed greater improvement than the uncoached students.

Beyond these common elements, the studies varied considerably in the coaching materials used, the number of hours of coaching the students had, and the persons who did the coaching (high school teachers in some cases and teachers at coaching schools in others). Some studies involved boys and girls in public schools; others only boys or only girls in private schools. The studies also differed in the ability level of the students involved and the courses they followed.

Changes in a student's Scholastic Aptitude Test scores between junior and senior years

Because no measure of scholastic ability is absolutely precise or reliable, a student's scores will change up or down from one time he takes the SAT to another time. Moreover, when several months elapse between the first and second times a student takes the test, he will tend to score higher the second time as a result of maturing intellectually during that period and, for those students who have had little experience with this type of objective test, as a result of practice. But such changes occur in about the same degree and with the same frequency *whether or not a student is coached.*

It is extremely difficult to predict what a student's score change will be on the SAT from March or May of his junior year to December or January of his senior year. Score changes vary considerably from

person to person. The variation in the size and also in the direction of SAT score changes is attributable primarily to measurement error rather than to variation among the students in intellectual growth. Similar variation in score changes is found when students take the same, or equivalent, tests on successive days. All tests, and in fact all measurement processes, are subject to such variation caused by error of measurement. Over the past several years, the average change in SAT-verbal or SAT-mathematical scores has been an increase in the

Table 1. Chances in 100 of obtaining various SAT*-verbal or -mathematical scores in the senior year, given a particular junior-year* SAT *score.*[1]

Senior-year SAT *scores*	*Junior-year* SAT *scores*				
	300	*400*	*500*	*600*	*700*
750-800					14
700-749				3	34
650-699				17	(35)
600-649			5	36	(15)
550-599		1	21	(32)	(2)
500-549		7	39	(11)	
450-499	1	26	(27)	(1)	
400-449	11	39	(7)		
350-399	31	(22)	(1)		
300-349	37	(5)			
250-299	(17)				
200-249	(3)				
Average senior-year SAT *scores*	*340*	*430*	*520*	*605*	*695*

1. Based on approximately 975,000 candidates who took the SAT in March or May of the junior year and in December of their senior year in 1963, 1964, 1965, or 1966.

Note: Parentheses denote frequencies of senior-year scores that are lower than the corresponding junior-year scores.

neighborhood of 20 points. Larger score increases are typical for students with lower initial scores, and smaller increases are typical for students with higher initial scores. Students who have studied mathematics in their senior year obtain on the average slightly larger mathematical score increases between their junior and senior years than do students who have not studied mathematics courses as seniors.

Table 1 gives the distribution and the average of the senior-year SAT scores associated with selected junior-year SAT scores. For example, students whose SAT-verbal scores in their junior year are 400 obtain, on the average, SAT-verbal scores of approximately 430 in their senior year. Similarly, the average SAT-mathematical score of seniors whose junior-year SAT-mathematical scores were 600 is approximately 605.

Table 1 makes it clear that there is a strong relationship between junior-year SAT scores and senior-year SAT scores. In spite of this strong relationship, however, it is also seen that within any group of individuals who score at the same level in their junior year, there is a fairly wide spread of scores of those students when they come to take the SAT in their senior year. Some of these students earn much higher scores in their senior year than they do in their junior year. Some of them actually lose score points, with relatively more and greater losses occurring at the upper end of the scale than at the lower end. But on the average, students gain on the second testing.

The significance of the average gain, as well as the variability of the gains, is that the SAT scores tend to increase as a matter of course, and in the case of a few students in every hundred, they increase by 100 points or more from one testing to another, whether or not these students were coached. It is therefore clear that large increases in the scores of some students who have been coached do not represent proof that coaching has produced the increases. Dependable evidence about the effectiveness or ineffectiveness of coaching can be observed in the average score changes of large and comparable groups of coached and

uncoached students. Such average changes are the chief findings of the studies reported here.

Statement by College Board trustees

As a result of the studies summarized in this book, the trustees of the College Board have issued the following statement testifying that increases in scores on the SAT that may result from coaching are negligible.

"The trustees of the College Entrance Examination Board have noted with concern the increasing tendency of secondary school students to seek the assistance of special tutors or of special drill at school in the hope of improving thereby scores earned on College Board examinations. The Board has now completed four studies designed to evaluate the effect of special tutoring or 'coaching' upon the Scholastic Aptitude Test, the basic test offered. Three other studies have been conducted independently by public high schools. These studies being completed, we now feel able to make the following statement:

"The evidence collected leads us to conclude that intensive drill for the SAT, either on its verbal or its mathematical part, is at best likely to yield insignificant increases in scores. The magnitudes of the increases which have been found vary slightly from study to study, but they are always small and appear to be independent of the particular method of coaching used and of the level of ability of the students being coached. The results of the coaching studies which have thus far been completed indicate that average increases of less than 10 points on a 600 point scale can be expected. It is not reasonable to believe that admissions decisions can be affected by such small changes in scores. This is especially true since the tests are merely supplementary to the school record and other evidence taken

into account by admissions officers.

"The conclusion stated here has been reached slowly and with care, although the atmosphere in which the problem has been studied has not been entirely calm. In recent years newspapers and even radio advertisements advancing the claims of the drillmasters have increased in number and boldness. Parents, already disturbed by exaggerated notions of the difficulties of students in gaining admission to college, have demanded that the schools divert teaching energy and time to a kind of drill that is obnoxious to educators of every philosophy.

"With parental concern so great, each completed study yielding negative findings with regard to the usefulness of coaching has led only to speculation that under some other set of circumstances some other set of students might make important score increases as a result of coaching for the test. The time has come to say that we do not believe it.

"Tutors often show apparent good results mainly because students and scores *do* change with the passage of time. Our studies have simply shown that the scores of students who are left alone change in the same directions and to nearly the same degree as do scores of students who are tutored. The public, though, is disconcerted to see any change in a measure of 'aptitude' which is regarded as unchangeable. As the College Board uses the term, aptitude is not something fixed and impervious to influence by the way the child lives and is taught. Rather, this particular Scholastic Aptitude Test is a measure of abilities that seem to grow slowly and stubbornly, profoundly influenced by conditions at home and at school over the years, but not responding to hasty attempts to relive a young lifetime.

"In addition to changes due to growth, other changes occur because the test, while dependable, shares a characteristic common to all tests in that it cannot be made to give exactly the same score for

each student each time the test is taken. Changes due to this lack of complete dependability are uncontrollable. Thus, with scores being affected by both the imperfect nature of the testing process and the student's growth, about one student in four will find that his scores actually decrease from one year to the next, while most other students will have small to moderate increases. About one student in fifteen will find that his scores increase by 100 points or more between junior and senior years in high school, *and this is true whether he is coached or not.* It is not surprising then that tutors are often able to point to particular students who have made large increases in their scores.

"It is possible to predict the size and number of fluctuations in scores that will occur within large groups, but fluctuations of individual scores cannot be predicted. Yet it is upon individuals that interest properly focuses, so that unexpected changes are easily, though erroneously, attributed to coaching, to the school, or to some other visible agency.

"We have said nothing about the tests of achievement in specific school subjects. These have not been studied in the same way as has the aptitude test. We do know that these tests do a modest but useful job of measuring learning of the material tested. We suspect that the question of coaching for these tests is a matter of choosing a method of teaching the subject. We cannot believe that drill on sample test questions is the most productive method available.

"Finally, we worry very little when parents of comfortable means decide that at worst tutoring can do no harm and therefore use their money for coaching toward College Board examinations. We are very concerned when parents purchase coaching they cannot afford or, failing to do so, feel that an unfair advantage has gone to those who have had a few weeks or months of tutoring. But we are concerned most, and have been moved to make this statement, because we see

the educational process unwillingly corrupted in some schools to gain ends which we believe to be not only unworthy but, ironically, unattainable."[1]

1. This statement was prepared by the trustees of the College Board for the information and advice of schools and colleges that had expressed concern about test coaching. It may be reprinted by schools and colleges in any format they may find convenient.

Seven studies of coaching

Coaching boys of above-average ability in independent preparatory schools (Dyer, 1953)

In 1953, Henry S. Dyer reported on the first of the four coaching studies sponsored by the College Board. He was then an associate director of the College Board and is now a vice president of Educational Testing Service.

Two independent preparatory schools for boys, School X and School Y, cooperated in the experiment. The schools were selected so that all students at both schools would be comparable in background and ability and in the kinds of instruction they would be given in their senior year. In both schools, instruction "was of a superior order," according to Dyer, and "classes were small and standards of performance high." There were 239 boys in School X and 229 in School Y who took part in the experiment. The seniors in both schools were given the SAT in September 1951 and again in March 1952. In the time between September and March, the boys in School X were coached for an average of six hours on verbal material and an average of four hours on mathematical material. During this period of coaching, the seniors at School Y were not coached in any way and had no special assignments in addition to their usual schoolwork.

The following factors were controlled for this study: the number of years each student had been enrolled in his school before September 1951; the number and level of foreign-language courses taken during senior year; the number and level of mathematics courses taken during senior year; the student's initial SAT-verbal and SAT-mathematical scores, earned in September.

School X coached its seniors from November through the first week in March. English teachers did the coaching for the verbal sections of the test, using a series of practice exercises consisting of questions

similar to those on the SAT itself. At regular intervals students worked an exercise and held class discussion of it for 10 to 30 minutes. Mathematics teachers coached the boys for the mathematical sections with practice questions similar to those on the mathematical sections of the test. Near the end of the coaching, the School X boys were given two special 40-minute versions of the SAT, one consisting entirely of verbal questions and the other entirely of mathematical questions. Each version was scored and then discussed in class.

One important feature of all the practice exercises, except the final 40-minute exercises, was that neither instructors nor students were given answers to the questions. Instead, all involved had to identify the right answers by agreement and argument. Answer keys were withheld by design. Dyer explained: "It was supposed that if all concerned fixed their attention on the *appropriate processes* rather than the *right answers*, the drill would have maximum effect."

The results of the study were analyzed by the statistical procedure known as "analysis of covariance." This procedure makes it possible to adjust all controlled factors simultaneously. In effect, then, each student in School X could be matched evenly with a student in School Y.

It was found that the average SAT-verbal score increase was 4.6 points higher for the coached boys in School X than for the uncoached boys in the School Y control group. This difference was not "statistically significant"—that is, it was not larger than could normally have been expected to occur by sheer chance.

The average SAT-mathematical score increase was 12.9 points higher for the coached boys in School X than for the uncoached boys in School Y. This difference was statistically significant: it was larger than would have been expected to occur by chance. A further breakdown of this result, however, showed the following: School X boys who had taken no mathematics courses as seniors but had been

given coaching in mathematics averaged 29.2 points higher in SAT-mathematical score increase than School Y boys who had taken no mathematics as seniors. By contrast, boys in School X who had taken mathematics as seniors averaged only 3.3 points higher in SAT-mathematical score increase than boys in School Y who had taken mathematics as seniors—despite the fact that the School X boys had been coached and the School Y boys had not.

Dyer concluded that "...coaching of able students in the senior year of a good secondary school:

"1. Is not likely to improve the SAT-verbal score by any amount.

"2. Will be of some advantage in raising the SAT-mathematical score only if the students happen not to be enrolled in regular mathematics courses."

Coaching boys and girls in public schools (French, 1955)

To investigate other questions not directly answered in Dyer's study, John W. French carried out two experiments on which he reported in 1955. French was a senior research psychologist in the research division at Educational Testing Service, and his studies were sponsored by the College Board.

In his larger study he looked into the question of the effect of coaching on public school students, both boys and girls. In the smaller study, which was an extension of the larger one, he explored the effect of coaching in an extreme case: coaching with the same questions that were to appear in the edition of the SAT the students would take after coaching.

Students and teachers at three public high schools took part in the studies. School A is in Michigan, and Schools B and C are in Massachusetts. The students who participated in the studies were the seniors at these schools who planned to go to college and a few stu-

dents taking technical courses at School C. All were given the SAT in September 1954.

At School A only the regular courses were pursued, and no attempt was made to coach the students for the SAT. Students at this school provided the experimental control group for studies of both the SAT-verbal and SAT-mathematical score changes.

At school B, the 126 students involved in the study had the following coaching on verbal materials in the winter semester: an average of 90 minutes on the school's special 500-word vocabulary list; an average of 90 minutes with vocabulary and reading comprehension materials in a commercially published English review book; an average of 90 minutes with sample verbal questions like those used in the SAT. Of these students, there were 44 who were planning to take the SAT in connection with their applications for college admission. These 44 undertook additional study of about two hours with the English review book and about two hours with three commercial publications on how to prepare for college entrance tests. School B students received no extra instruction or coaching in mathematics.

At School C the 188 college preparatory and technical students participating in the study were coached throughout the winter semester with both verbal and mathematical materials. These materials consisted of 20 special exercises made up of items specially prepared at Educational Testing Service to resemble items on the SAT. The exercises were worked in class and discussed immediately. There were 39 students who expected to take the SAT for college admission, and these 39 were given extra drill on mathematics problems. Shortly before the March 1955 administration of the SAT all 188 students took a 40-minute practice test of verbal items and a 40-minute practice test of mathematical items.

In March 1955 the students in all three schools took the SAT again. It was found that the average scores for various groups of the

coached students had increased from September to March more than the average scores of the corresponding groups of uncoached students. The differences in score gains were comparatively small, however, as shown in Table 2.

Table 2. Differences in average SAT *score gains of coached and uncoached students*

	Average score-gain differences	
Verbal	*School C minus School A*	*School B minus School A*
Boys	15.2 (.02)	–1.7
Girls	22.4 (.01)	12.3
Boys and girls	18.3 (.01)	5.0 (.05)
	Average score-gain differences	
Mathematical	*School C minus School A*	*School C minus School B*
Boys not taking mathematics	9.3	29.2 (.05)
Girls not taking mathematics	.9	4.2
Boys taking mathematics	5.0	18.8 (.05)
Girls taking mathematics	20.2	29.7 (.05)
All boys and girls	6.2 (.01)	18.0 (.01)

Notes: Students at School A were not coached. Students at School B were coached in verbal material only. Students at School C were coached in both verbal and mathematical material.

Differences in average score gains are expressed in points on the College Board score scale of 200-800. Positive numbers show differences in favor of the coached groups. The two-place decimals in parentheses after some of the gains indicate the probability that those gains could have occurred by chance. The decimal .02, for example, may be read as "there are no more than 2 chances in 100 that the foregoing result could have occurred by chance." Differences not followed by such decimals could have occurred by chance more than 5 times out of 100 and are therefore not considered statistically significant.

The main conclusion from this study is that the largest average gains attributable to coaching, for groups that included both boys and girls, were only about 18 points for both the SAT-verbal and the SAT-mathematical score. A score difference of 18 points is ordinarily of little or no practical importance. On the SAT score scale of 600 points, it is equivalent to a gain of slightly more than 1 point on the traditional 60 to 100 classroom grading scale. Moreover, it is smaller than the "standard errors of measurement" of the SAT. The standard error of measurement of a test is an indication of the variability of a score on the test. For example, the standard error of measurement for the SAT-verbal score is about 32 points; this means that there is about 1 chance in 3 that a student's SAT-verbal score will vary 32 points or more either up or down from his theoretical true score. The 18-point gain found in this study to result from coaching is well below this 32-point standard error. The same conclusion applies to the SAT-mathematical score, which has a standard error of about 34 points.

The study of public school students agreed with Dyer's study of independent school students in its most important findings: The difference in scores resulting from coaching is not large in comparison to the standard error of the test. This conclusion holds true even for the girls in the study, who tended to gain more through coaching than the boys, and it holds true also for intensive coaching.

In contrast to Dyer's findings in his independent school study, French did not find that coaching led to consistently larger average gains in SAT-mathematical scores by students who were not taking mathematics as seniors.

In reporting on this study, French concluded: "Specific coaching for the SAT does raise the test scores slightly, on the average, but so slightly that the same amount of study on regular school assignments is likely to help just as much on the test and much more toward the student's understanding of the subject."

Coaching students by using actual items from the SAT *they later take (French, 1955)*

In the smaller study that grew out of the larger one, French looked into the question of the effect of coaching students on the same questions that appear on the SAT when they take it the second time. He called this the "identical item" study.

Half of the questions on the two 40-minute final practice tests given the School C students in French's study were the same as questions in coaching exercises the students worked with previously. Scores made by the students on these identical items in the practice tests were computed separately. The identical-item scores were then converted to the equivalents of scores on the full SAT. It was found that coaching on identical items produced an average increase 47 points higher on the verbal score and 15 points higher on the mathematical score than the average increases produced by coaching without the benefit of identical items.

This inquiry was made because it is believed that the operators of some coaching programs obtain some items that have been included on editions of the SAT used in past administrations of the test. Since some items are used in more than one edition of the SAT, students coached with items from past editions of the test may find a few of these items on a current form. The results of the study indicate that even if it were possible to coach with as many as one-tenth of the items that would be used on a future form of the SAT—a virtual impossibility—the effect on the scores of students coached with these items would be negligible.

Intensive coaching of small groups of students (Dear, 1958)

One question left unanswered after the studies by Dyer and French was that of the effect of intensive coaching in small, tutorial groups.

Moderate amounts of coaching in classes result in SAT score increases too small to be of any practical importance, but what happens if students get many hours of coaching, given with close personal attention?

This was the main question examined in a fourth coaching study sponsored by the College Board. It was made in 1956 and 1957 by Robert E. Dear, a research psychologist in the research division of ETS. His report was issued in 1958.

Nineteen secondary schools in New Jersey and in the New York City and Philadelphia areas took part in the study. Selected students in 10 of the 19 schools were coached; the 10 consisted of six public schools, two independent schools for boys, and two independent schools for girls. Students at the other nine schools provided a control group in the experiment; five of these schools were public and four were independent. All 19 of the schools had been picked at random from schools in which there were at least 15 junior-year students who had taken the SAT in May 1956.

Six students in each of the first 10 schools were picked to receive coaching. The students were selected from those who wanted to take part, and all were seniors who had taken the SAT as juniors in May 1956. Each school's six included two students whose May 1956 SAT-verbal and SAT-mathematical scores combined totaled 1,100 or more, two whose combined scores totaled 900 to 1,100, and two whose scores totaled below 900. Students who participated in the study were chosen at random from within these categories. For each set of two students, a student who had similar qualifications served as an alternate in case one of the two should be unable to complete the experiment. Altogether, 71 students were coached—practically all the original group and about a third of the alternates. In the 10 schools that gave coaching, there were comparable students who did not receive coaching, and these students were used as one control

group. Comparable students in the other nine schools that did not give coaching served as a second control group.

Two tutors were picked in each of the 10 schools to do the coaching, one to coach with verbal material and the other with mathematical material. The tutors were teachers at the schools who were recommended by the heads of the schools and were paid for the extra work.

The tutors were provided 20 sets of verbal practice exercises and 10 sets of mathematical practice exercises prepared by ETS to be similar to the SAT. Also supplied by ETS were four practice tests (two verbal and two mathematical) and copies of the College Board booklet, *A Description of the College Board Scholastic Aptitude Test*, which describes and gives many sample questions from the test. Working these practice exercises and analyzing the results made up a large part of the coaching. Tutors were instructed to use whatever methods seemed best for the individual student, and they met midway in the experiment for a one-day workshop to exchange ideas.

Coaching began in mid-November 1956. Students were coached in pairs (or threes in some cases, if alternates joined the subjects), and each pair studied with the verbal tutor one class period a week and with the mathematical tutor one period a week. The students were also given about an hour's homework every week, on verbal material one week and mathematical material the next. Each student thus spent about three hours a week on coaching. In mid-January, 60 of the coached students took the SAT a second time, to meet college admissions requirements. By then, they had spent about 18 hours in six weeks on coaching. The students continued with coaching for the three hours a week into early March, so that they each had about 36 hours of coaching in 12 weeks before taking the SAT again in March.

The results of this study were analyzed by means of analysis of covariance, so that adjustments could be made for differences of ability among the experimental and control groups. The 60 coached

students who took the SAT in January 1957 showed an average SAT-verbal gain over their May 1956 scores that was 2.91 points less than the corresponding average gain of the uncoached students in the same schools and 2.13 points less than the average gain of uncoached students in the other nine schools.

For those who took the SAT in March, a similar analysis of gains in adjusted mean verbal scores was not usable, but the unadjusted mean difference was approximately 6 points higher for the coached group than for either control group—a difference of no practical importance.

The coached group showed an average gain of about 26 points more than the average gain of the pooled control groups on the March 1957 SAT-mathematical scores. Similarly, coached students showed a larger average gain of about 21 points on the January 1957 SAT-mathematical scores, compared to the uncoached students of the two control groups.

Dear concluded that "these average gains in mathematical scores, though statistically significant, are probably not of sufficient magnitude to give coached students an unfair advantage in college admissions." And it might be noted that the average gains of 26 and 21 points are less than the SAT-mathematical standard error of measurement, which is about 34 points.

Commercial coaching programs: their effect on public-school boys and girls of above-average ability (Frankel, 1960)

In 1959, Edward Frankel made a study of the effect of coaching in private commercial courses outside school. Frankel was then senior guidance counselor at The Bronx High School of Science, which sponsored this study. He is now an associate professor of education at Hunter College.

Frankel selected 45 matched pairs of second-semester seniors in the spring of 1959. They were matched so that:

1. Each pair consisted of a student who had taken a commercial coaching course and one who had not.

2. Both members of the pair had taken the SAT in May 1958 and had scored within 10 points of one another on both the verbal and mathematical parts of the test.

3. The "coached" member of the pair had taken a coaching course between May 1958 and December 1958 or January 1959.

4. Both members of the pair had taken the same edition of the SAT as seniors—that is, both students in 24 of the pairs had taken the December 1958 SAT, and both students in the other 21 pairs had taken the January 1959 SAT.

5. Boys were matched with boys and girls with girls.

The average SAT scores obtained in May 1958 by members of the two groups were approximately equal: average SAT-verbal scores were 560.58 for the coached group and 559.20 for the uncoached group; average SAT-mathematical scores were 575.55 for the coached group and 573.44 for the uncoached group.

The 45 students who had been in commercial coaching programs reported that they had attended in groups of approximately 25 students and had been coached in 10 sessions of three hours each for a total of 30 coaching hours. The coaching had taken place during their senior year just before the December or January SAT.

Frankel's main finding was that the average increase in the SAT-verbal score was 8.38 points greater for the coached group than for the uncoached group and that the average increase in the SAT-mathematical score was 9.07 points greater for the coached group than for the uncoached group. Neither difference was statistically significant.

Within the coached group, changes in SAT-verbal scores from the first to the second testing ranged from a loss of 46 points to a gain

of 147 points; changes in SAT-mathematical scores similarly ranged from a loss of 29 points to a gain of 218 points. Among the uncoached students, changes in SAT-verbal scores ranged from a loss of 80 points to a gain of 199 points; changes in SAT-mathematical scores ranged from a loss of 19 points to a gain of 203 points.

Frankel concluded that for the group under study, the coaching had produced no significant gain in either the SAT-verbal or the SAT-mathematical score.

Intensive tutoring of public-school boys and girls in vocabulary, reading skills, concept mastery, problem solving (Whitla, 1962)

Another kind of coaching outside school was studied in 1959 and 1960 by Dean K. Whitla, assistant professor of education and director of the Office of Tests at Harvard University. The study was conducted with the cooperation of the Reading Institute of Boston, a proprietary school organized to help students develop study habits and reading skills.

Fifty-two students who were enrolled in the Reading Institute in the fall of 1959 and later completed its tutoring program served as the subjects of the study. They came from a variety of secondary schools in the Boston area, primarily public schools. All 52 had taken the SAT in the previous March or May as juniors, and they were tested with another edition of the SAT when they began their Reading Institute work.

Whitla described the tutoring program as follows: "In its course of five two-hour sessions and intensive homework drill, the Reading Institute seeks to affect just that set of skills which the Scholastic Aptitude Test is intended to measure: an ability to understand and organize verbal concepts and to solve problems in quantitative and verbal areas. The class sessions are very intensive, covering vocabu-

lary, reading skills, mathematical concepts, and experience in dealing with multiple-choice items. The homework assignments include vocabulary lessons, mathematical problems and extensive reading."

A second group of 52 students served as a control group. They were selected from the Newton, Revere, and Watertown high schools, which are public schools in large Boston suburbs and are typical of the schools attended by the 52 students in the Reading Institute programs. These students, too, had taken the SAT as juniors, and they took the alternate form of the SAT used as a "pretest" at the beginning of the study.

Both the coached and the uncoached students took the SAT again in January 1960. The average scores obtained by the two groups in their junior year and on the two administrations in their senior year are given in Table 3.

Table 3. Average SAT *scores of coached and uncoached students*

	Junior year (March or May)		*Pretest*		*Senior year (January)*	
	Coached	*Uncoached*	*Coached*	*Uncoached*	*Coached*	*Uncoached*
Verbal	413.27	412.50	431.35	431.35	462.12	451.15
Mathematical	437.20	437.20	467.88	469.60	468.32	475.32

Although the coached group's average SAT-verbal score at the end of the study was about 11 points higher than the average SAT-verbal score of the uncoached group, and its average SAT-mathematical score was approximately 7 points lower than that of the uncoached group, neither of these differences is statistically significant.

Whitla also reported on the frequencies with which relatively large score increases and decreases occurred from junior year to senior year among the coached and uncoached groups. The numbers of students whose score changes were relatively large are given in Table 4.

Table 4. Numbers of students whose scores changed 50 or more points between junior year and senior year

	50-99 points		100 or more points	
SAT-verbal score (52 pairs of students)	Coached	Uncoached	Coached	Uncoached
Increase	16	16	6	3
Decrease	1			
SAT-mathematical score (50 pairs of students)				
Increase	14	18	7	5
Decrease	6	3		

Although there were six coached students who increased their SAT-verbal scores by 100 or more points and only three uncoached students, there were no more of these large increases than could have occurred by chance. With respect to the numbers of students whose SAT-mathematical scores increased, the uncoached students obtained a greater number of increases and fewer decreases than did the coached students, but these differences between the two groups were not statistically significant.

Whitla concluded: "What we have examined is whether training beyond that offered in suburban Boston high schools of both high and low socioeconomic communities contributes to scoring on the Scholastic Aptitude Test. We have found that it does not."

Informal review of test scores of coached and uncoached students within one public school (Lass, 1958)

An informal study of the effect of coaching among students at the Abraham Lincoln High School in Brooklyn, New York, was made in 1958 by Abraham H. Lass, principal of the school. More than three-

fourths of the graduates of Abraham Lincoln High School go on to college. Lass made his study to obtain information on which the school could base advice to parents and students about coaching for the SAT.

Three groups of students in the school's senior class of 1957-58 were considered in the study. All had taken the SAT as juniors in the spring of 1957 and took it again as seniors. One group consisted of 82 students who had had no coaching or other special preparation for the test. Another group of 38 students had taken coaching in programs outside the school before taking the SAT as seniors. A third group of 21 students had attended an orientation program for the SAT given by the school. The orientation program, presented by the school's teachers, consisted largely of acquainting students with the way the SAT is administered and the kinds of questions it contains. The program did not involve extensive drills on multiple-choice questions or other typical coaching activities. The average SAT scores earned by the three groups in their junior and senior years and the average score gains for each group are given in Table 5.

Table 5. Average score gains of three groups of students at Abraham Lincoln High School

SAT-*verbal score*	*Junior year*	*Senior year*	*Average gain*
Uncoached students	494	535	41
Coached students	462	506	44
Students who took orientation course	499	552	53
SAT-*mathematical score*			
Uncoached students	517	570	53
Coached students	475	539	64
Students who took orientation course	506	558	52

The results shown in Table 5 do not necessarily apply to any other students or schools, primarily because the study provided no control over differences in ability from one group to another. In particular, the average scores of the coached group were lower than the averages for the other two groups in the junior year. Nevertheless, for these students it is clear that coaching in courses outside school produced no advantage in the case of the verbal score and, at best, an inconsequential advantage in the case of the mathematical score.

Sources

Dear, Robert E., *The Effects of a Program of Intensive Coaching on* SAT *Scores.* College Entrance Examination Board Research and Development Reports, RB-58-5. Princeton, N.J.: Educational Testing Service, 1958.

Dyer, Henry S., "Does Coaching Help?" *College Board Review,* No. 19, February 1953, pp. 331-335.

Frankel, Edward, "Effects of Growth, Practice, and Coaching on Scholastic Aptitude Test Scores." *Personnel and Guidance Journal,* May 1960, Vol. 38, No. 9, pp. 713-719.

French, John W., "An Answer to Test Coaching," *College Board Review,* No. 27, Fall 1955, pp. 5-7.

Lass, Abraham H., Unpublished Report. Brooklyn, N.Y.: Abraham Lincoln High School.

Whitla, Dean K., "Effect of Tutoring on Scholastic Aptitude Test Scores." *Personnel and Guidance Journal,* September 1962, Vol. 41, No. 1, pp. 32-37.